California Cancer Registry Automation

White Paper on Improving Data Timeliness

September 2003

To D+6… and Beyond to Real-Time Reporting

(DIAGNOSIS PLUS SIX MONTHS)

D+6, the practice of creating research quality data six months following patient diagnosis of cancer, is an initiative the CCR should consider pursuing to support the research mission of producing timely, complete and high quality data.

As the market for research funds is ever changing, the management of "research data" and its associated risk becomes ever more important. D+6 is a key initiative to reduce risk inherent in producing cancer research related data. The CCR will realize increasing efficiencies and reduced failures with a move to D+6. There are many drivers for D+6: organizational irrelevance at D+20 of producing "old data", increasing case volume, contractual obligations, cost savings, and research funding competitiveness. CCR participants (all CCR organizations) must radically change their systems and procedures in order to make "D+6…. and beyond" feasible. First and foremost, they must abandon a mentality of batch processing and think in terms of continuous, real-time processes. They must also embrace an exception-based orientation to their workload eliminating manual processes. Technology will be vital in helping CCR participants succeed in this endeavor. Key elements of the new processing environment include event-driven processing, business rules management, electronic interfaces, straight-thru-processing, data warehouses, real-time tracking of deliverables, and continuous quality control.

Why D+6?

What is D+6?

D+6, a possible major initiative in support of the CCR Mission, refers to the practice of making data available for research six months following patient diagnosis of cancer. Currently, diagnosis patient set information generally takes 20 months before being "research-ready". Shortening this interval will bring benefits both in efficiency and research opportunities.

The fact that there is a significant delay between cancer data creation and data that is "research-ready" becomes increasingly unnecessary in a world of fast electronic communications. By contrast, a delay between data creation and "research ready" data is unavoidable in an environment of physical handling and manual review and correction. Additionally, many of today's reporting sources observe a multi-month (six) reporting convention that is based on this model of iterative exchange of physical paper. In reality, with modern electronic communications and converged databases, reporting could happen almost at the same time as the data creation. D+6 is a step in the direction of instant reporting.

Drivers for D+6

Delays pose risk to organizational relevance
The long delay between cancer diagnosis data and research-ready data poses high risks of making CCR an anachronism. For example, the population of a fast growing community can change significantly between the date a cancer is actually discovered and later published by CCR. The shorter delay between cancer discovery and research, the lower the risk of CCR becoming irrelevant. In the ideal situation, cancer discovery and the complete research-ready reporting of the cancer would happen almost simultaneously.

Increasing case volume
As the State of California population grows and ages, the reported number of cancer cases increases. During the complete process overhaul that is necessary to enable D+6, the CCR has the opportunity to improve processes so that larger volumes can be handled more productively. More timely, more complete, and less costly data is possible with updated processes using modern technology.

Contractual obligations
All CCR participants are under contractual obligations to produce data. Most have multiple contractual obligations to various funding sources. D+6 greatly reduces the risk of being out-of-contract because of "late" data production. This is similar to the "six-sigma" approach adopted by the manufacturing industry. By shortening the data production life cycle, it provides "breathing-room" for errors, mistakes and unanticipated problems.

Cost savings
There will be significant cost savings as a result of the move to "D+6… and beyond". Per "data-unit" cost savings of 50% and greater are conceivable. Such savings would come from fewer errors and corrections, and reduced manual intervention. For example: by implementing "exception-based" technologies it is possible to reduce manual intervention by 80% in Quality Control (Visual Editing), Linkage/Matching and Consolidation of patient set information. Further use of this technology in Updates/Corrections, Abstracting, and Follow-Up would produce additional cost savings.

Research competitiveness
Another important factor driving the D+6 initiative is the importance CCR places on its reputation as the data source of choice for the cancer research community. Researchers often have a choice of data sources. Funders also have a choice where to invest. If the CCR does not continue to innovate, streamline processes, increase timeliness, improve data quality and data completeness, the research and funding community could look elsewhere. Other data sources will sooner or later modernize toward this goal.

What will change?

The change to "D+6… and beyond" requires a complete paradigm shift across the CCR and participant organizations. The focus will necessarily change to real-time and continuous 24-hour processing of cancer related data and away from batch processing. Cancer data gathering personnel will have real-time and on-line access to information. They will be able to resolve problems quickly, and see the results of their fixes instantaneously. Problems will be identified and fixed at the "root cause". Extensive use of exception-based and other automation tools and systems will be used.

Straight-thru-processing (STP)
The CCR should consider embracing "straight-through-processing" (STP) which aims for abstract to research-ready processing with no human intervention, when possible. The CCR participants would aim for a high percentage of cases – those that are straightforward and routine – to be completed without being reviewed by humans. This would free up operations staff at each participant organization to focus on resolving exceptions, or problems. The CCR could target automating processes like Abstracting, Visual Editing, Linkage, Follow-Up, and Consolidation so that routine cases can be processed with little or no human intervention.

Ideal Abstracting
Ideal Abstracting works on the principle that the earlier you discover a problem, the easier it is to fix. The Ideal Abstract will be error free and timely. Any discrepancies are resolved by the abstractor who actually prepared the case, rather than by QC staff waiting days later to review and resolve problems. Currently, each side reports facts somewhat independently. It is often weeks/months before discrepancies come to light. By this time, the original party may have forgotten many details surrounding the abstract. Developing intelligence to create "ideal abstracts" straight away will make reporting more accurate and efficient. A related long term initiative would be to partner with suppliers of hospital software systems and create a "seamless" transfer of cancer related data into the abstract with little or no human intervention.

Real-time cancer data reporting
While cancer discovery happens continuously throughout the day (and all night as well), discoveries are not reported as they happen. Many reporting sources wait until late in the reporting cycle (if at all) to report information, often in bulk. This puts strain on the internal systems and resources of the regional registries as well as the centralized systems and resources at CCR. Since processing these late-reported cases goes well beyond the reporting cycle, both reporting sources and regional registries often do not meet delivery obligations. In a "real time" reporting environment, reporting happens well within the reporting cycle and allows all participants time to resolve discrepancies.

Expanded reporting details
Currently some reporting sources do not report all cancer-related information to the CCR participants. Instead, they submit partial information. This makes it difficult for the CCR participants to accurately edit, link and consolidate patient set information. Reporting sources will need to abandon this practice, reporting all cancer related information in its entirety as it happens.

Standard messaging
In the fast-paced D+6 environment, quick, accurate communication is essential. The cancer data collection and research organizations need to agree on standards for communicating among all participants (Pathology Laboratories, Department of Motor Vehicles, Vital Records, Environmental Health, etc.). Currently there is a patchwork quilt of interfaces between parties: some proprietary electronic, some standards-based electronic, some manual phone or fax, and some manual Internet (i.e. electronic mail). Agreement on a small number of non-proprietary standards for communicating will greatly help the cancer research community achieve D+6.

Non research-ready data uses
Non research-ready data have long been accorded second class status. CCR participants have viewed "in-processing" data as useful only to operations departments. With current technology, timely and reliable "in-processing" data can become another asset to be managed as both a research and production opportunity. The D+6 initiative will transform previously "under-valued" data from

the back-room to the forefront as another competitive research asset.

How can technology help?

CCR participants have spent significant amounts of money on updating "legacy" type systems. Additionally, in the latter half of the 1990s, some of the system maintenance efforts were concentrated on Y2K. But, many of the core business processes and systems have not undergone significant change in many areas, in that *manual processes still abound.*

For D+6 to succeed it is imperative that business processes are updated so that manual work is eliminated where possible; and that CCR explore technological solutions within the walls of CCR, between internal participants, and with external reporting participants.

Exception-based orientation
Movement towards "exception-based" systems is critical. Exception-based systems process the bulk of patient related information without human intervention, leaving only the problem items for humans to fix. User interfaces highlight incomplete, non-standard, failed, complex or other exception conditions. These exceptions are routed to the correct operations staff as they arise. This system saves staff from having to perform routine and mundane steps that do not add value to data quality. As a problem arises, it is directed to the appropriate work queue and can be resolved on the same day. Exception-based systems efficiently utilize scarce resources by allowing the computers to do what they excel at – routine, repetitive tasks – freeing human staff to solve more complex problems that require their operations knowledge.
Well designed exception-based systems push problems directly to the person who will resolve them. By detecting problems early in the data input cycle, there is a good chance that the person who created the problem will be the one to resolve it. In the example of new case processing all abstraction errors are immediately directed back to the relevant abstractor for review. This is different from the current model where QC operations staff resolve the issue days/months later, often with incomplete information.

Electronic interfaces
Electronic interfaces are a vast improvement over phone, paper, and fax communication, which are still used by CCR participants in many situations. This is because electronic interfaces mandate data exchange in a standard form, and because they operate directly between the computer systems of participants. Interfaces can be either between peers or between organizations performing different functions. The move is on to standardize interface formats and expand the use of electronic interfaces. National, Federal, and State organizations are looking at ways to eliminate proprietary links and create new, generalized, and expandable interfaces based on updated technologies. Electronic interfaces between all Pathology Laboratories and CCR is a practical example of this solution that would have significant positive impact.

Business Rules Management
Movement from interpretable paper-based standards used in manual decision making, towards structured "actionable" business rules that can be programmed to produce consistent and predictable results offers the possibility of executing policies and procedures without misinterpretation. Deploying Business Rules Management (BRM) technology should be given serious consideration. It is possible that a BRM exception-based system could process the bulk of patient related information without human intervention, leaving only the complex problem items for humans to resolve. Deployment of BRM technology could jump-start the movement to an

automated exception-based data processing environment.

Event-driven processing

Currently, most computer systems in the cancer data management industry use batch processing. For example, when Death Status is applied, an operations programmer loads a "batch file" of information into the database. Upon upload of the file, the following day personnel can view the results of the uploaded Death Status information.

In contrast, event-driven processing refers to computer processes that are triggered by events in the outside world. In the Death Status example, when the Death Status information is created at its source, the action triggers a process to update the CCR and other affected databases. In this way, the data in the CCR's databases reflect the situation in the real world.

Real-time tracking(metrics) of delivery obligations

CCR participants have a complex set of deliverables to track. Often, problems in meeting contract deliverables are discovered too late for corrective actions to have an effect. Real-time tracking of deliverable obligations would aide in identifying problems early in the data production cycle so that corrective actions would have a high rate of success. In such an environment, each activity results in the immediate and explicit capture of information as it relates to a pending deliverable obligation. At all points in time, the CCR participants are able to view progress in meeting deliverables. The system could be further developed to automatically alert the participant when deliverables are at risk of being missed.

Continuous Quality Controls (CQC)

Quality control currently happens late in the data quality cycle. In a real-time "CQC" environment, the system continuously monitors its data to ensure both internal system consistency and agreement with the "outside" (DS&QC) world. These checks are crucial - if the data in the production system are inaccurate, the system becomes less valuable. Internally, the system is designed to continuously reconcile various "audit-points". Ensuring that these "audit-points" are in agreement helps ensure accuracy of the system's data.

Internal system reconciliation does little good without checks from the "outside" external world. The "CQC" system also checks its positions versus external standards and audits. The "CQC" system would be designed to work in conjunction with external auditing to reconcile internal positions in real time with DS&QC staff.